Amendments to the Specification:

Please replace the paragraph beginning at page 1, line 6 as with the following amended paragraph:

This application is a divisional (and claims the benefit of priority under 35 USC 120) of United States Patent Application No. 09/502,426, filed February 11, 2000, which claims the benefit of U.S. Provisional Application ~~This application is related to provisional patent applications serial~~ Nos. 60/119,657, filed February 11, 1999 and 60/119,658, filed February 11, 1999~~, from which priority is claimed~~ under 35 USC § 119(e)(1), ~~and~~ which applications are incorporated herein by reference in their entireties.

Please replace the paragraph beginning at page 11, line 31 as with the following amended paragraph:

Figure 3 depicts alignment of cytochrome P450 proteins that exhibited the most similarity to DWF4 (SEQ. ID NO:2) in BLAST searches. GenBank accession numbers are AF044216 (DWF4; CYP90B) (SEQ. ID NO:2), X87368 (CPD; CYP90A), U54770 (tomato; CYP85), D64003 (cyanobacteria; CYP120), U32579 (maize; CYP88), U68234 (zebrafish; CYP26), and M13785 (human; CYP3A3X). Dashes indicate gaps introduced to maximize alignment. Domains indicated in Figure 2B are highlighted in a box. Amino acid residues that are conserved >50% between the compared sequences are highlighted by a reverse font, and identical residues between DWF4 and CPD are boxed and italicized. Open triangles are placed under the 100% conserved residues. Closed triangles locate functionally important amino acid residues, for example, threonine (T) at 369, which is thought to bind molecular oxygen, and cysteine (C) at 516, which links to a heme prosthetic group by a thiolate bond. X's indicate mutated residues in *dwf4* alleles. Multiple sequence alignment was performed using PILEUP in the Genetics Computer Group package, and box shading was made possible by the ALSCRIPT package (Barton (1993) *Protein Eng.* 6:37-40).

Please replace the paragraph beginning at page 17, line 1 as with the following amended paragraph:

Techniques for determining nucleic acid and amino acid "sequence identity" are known in the art. Typically, such techniques include determining the nucleotide sequence of the mRNA for a gene amino and/or determining the amino acid sequence encoded thereby, and comparing these sequences to a second nucleotide or amino acid sequence. In general, "identity" refers to an exact nucleotide-to-nucleotide or amino acid-to-amino acid correspondence of two polynucleotides or polypeptide sequences, respectively. Two or more sequences (polynucleotide or amino acid) can be compared by determining their "percent identity." The percent identity of two sequences, whether nucleic acid or amino acid sequences, is the number of exact matches between two aligned sequences divided by the length of the shorter sequences and multiplied by 100. An approximate alignment for nucleic acid sequences I provided by the local homology algorithm of Smith and Waterman, Advances in Applied Mathematics 2:482-489 (1981). This algorithm can be applied to amino acid sequences by using the scoring matrix developed by Dayhoff, Atlas of Protein Sequences and Structure, M.O. Dayhoff ed., 5 suppl. 3:353-358, National Biomedical Research Foundation, Washington , D.C., USA, and normalized by Gribskov, Nucl. Acids Res. 14(6):6745-6763 (1986). An exemplary implementation of this algorithm to determine percent identity of a sequence is provided by the Genetics Computer Group (Madison, WI) in the "BestFit" utility application. The default parameters for this method are described in the Wisconsin Sequence Analysis Package Program Manual, Version 8 (1995) (available from Genetics Computer Group, Madison, WI). A preferred method of establishing percent identity in the context of the present invention is to use the MPSRCH package of programs copyrighted by the University of Edinburgh, developed by John F. Collins and Shane S. Sturrok, and distributed by IntelliGenetics, Inc. (Mountain View, CA). From this suite of packages the Smith-Waterman algorithm can be employed where default parameters are used for the scoring table (for example, gap open penalty of 12, gap extension penalty of one, and a gap of six). From the data generated the "Match" value reflects "sequence identity." Other suitable

programs for calculating the percent identity or similarity between sequences are generally known in the art, for example, another alignment program is BLAST, used with default parameters. For example, BLASTN and BLASTP can be used using the following default parameters: genetic code = standard; filter = none; strand = both; cutoff = 60; expect = 10; Matrix = BLOSUM62; Descriptions = 50 sequences; sort by = HIGH SCORE; Databases = non-redundant, GenBank + EMBL + DDBJ +PDB + GenBank CDS translations + Swiss protein + Spupdate + PIR. ~~Details of these programs can be found at the following internet address: http://ncbi.nlm.gov/cgi-bin/BLAST.~~

Please replace the paragraph beginning at page 51, line 25 as with the following amended paragraph:

Analysis of the complete genomic sequence, starting at the EcoRI site, with the promoter prediction by neural network (NNPP) package (~~http://www.hgc.lbl.gove/projects/promoter.html~~), indicated that the gene included a putative promoter (TATAT is found in the putative promoter region between nucleotides –143 to –78) and polyadenylation signal sequences (AATAA near a position at 3238 bp and a putative GU-rich signature from 3283 to 3290 bp).

Please replace the paragraph beginning at page 55, line 28 as with the following amended paragraph:

Thus, phylogenetic analyses of these seven proteins with cytochrome P450s unique to plants (group A; Durst and Nelson (1995), *supra*) indicate that DWF4 does not cluster with these cytochrome P450s (Figure ~~6~~ 4). Rather, DWF4 clustered with cytochrome P450s from other organisms: cyanobacteria (CYP120), rat (CYP3A2), human (CYP3A3X), and plants (CYP90, CYP85, and CYP88). DWF4 also deviates from the consensus sequence in the group A heme binding domain in that it possesses a PFGGGPRLCAG sequence in which arginine (R) is substituted for proline (P). However, domain A of DWF4, AGHETS, fits the consensus of domain A of group A. These characteristics suggest that DWF4 is a monooxygenase, similar to P450s of group A, that utilizes molecular oxygen as a source of the hydroxyl group, but it

mediates some reaction(s) that are not necessarily specific for plants, for instance, steroid hormone biosynthesis, which is a critical event for animals. In fact, the similarity of DWF4 to the rat testosterone 6β-hydroxylase (34%; GenBank accession number 631895) or glucocorticoid-inducible hydroxylase (31%; Molowa et al. 1986; GenBank accession number M13785) supports this idea. Further, the similarity that DWF4 shares with CYP90A and CYP85, 66 and 59%, respectively, is additional proof that it is involved in plant steroid biosynthesis (Bishop et al. 1996; Szekeres et al. 1996).

Amend the paragraph beginning at page **54**, line **4**, as follows:

Durst and Nelson (1995) *Drug Metab. Drug Interact.* 12:189-206 classified plant cytochrome P450s into two distinct groups based on their clustering nature in a phylogenetic tree. All of the group A families cluster and are assumed to originate from a common plant P450 ancestor. The group A cytochrome P450s conform to the characteristic consensus sequences (A/G)GX(D/E)T(T/S) in domain A (also called helix I) and PFG(A/S/V)GRRXC(P/A/V)G (SEQ ID NO:26) of the heme binding domain (D) with only a few exceptions. Group A cytochrome P450s appear to catalyze plant-specific reactions such as lignin biosynthesis (Figure 6; GenBank accession number P48421). By contrast, P450s that do not belong to group A (non-A P450s) are scattered in the phylogenetic tree. They share more amino acid identity/similarity with P450s found in animals, microbes, and fungi than with those found in plants. The non-A P450s possess functions, such as steroid metabolism, that are not limited to plants. Generally, non-A P450s have limited homology with known domains described for group A.

Amend the paragraph beginning at page **55**, line **1**, as follows:

Six cytochrome P450 sequences with the greatest homology to DWF4, CYP90A1, CYP85, CYP88 (Winkler and Helentjaris (1995) *Plant Cell* 7:1307-1317; GenBank accession number U32579), cyanobacteria CYP120 (Kaneko et al. (1996) *DNA Res.* 3:109-136; GenBank accession number D64003), human CYP3A3X (Molowa et al. (1986) *Proc. Natl. Acad. Sci.* USA 83:5311-5315; GenBank accession number M13785), and zebrafish CYP26 (White et al.

White (1996) *J. Biol. Chem.* 271:29922-29927; GenBank accession number U68234), were chosen for multiple sequence alignment. Putative domains defined by Kalb and Loper (1988), *supra* are boxed and labeled in Figure 3. First, the heme binding domain pFGgFpRlCpGkel (SEQ ID NO:27) matches completely the sequence defined previously. Uppercase letters in the domain indicate amino acids conserved at all seven sequences in the alignment, and lower-case letters represent residues conserved in at least half of the proteins. Of the amino acids conserved in the heme binding domain, the function of the cysteinyl is established as a thiolate ligand to the heme (Poulos et al. (1985), *supra*).

Amend the paragraph beginning at page **55**, line **15**, as follows:

Domain A is defined by xllfaGhEttssxIxxa (SEQ ID NO:28). Lowercase x's indicate variable amino acids. An invariant glutamate (E) preceded threonine (T) at position 314, T314, which is believed to bind dioxygen, was conserved in all proteins compared except CYP88 of maize. The second signature sequence, domain B, is also conserved in *DWF4* with significant similarity. A valine at position 370 is conserved in all of the proteins, but is does not appear in Kalb and Loper's classic report (1988) on conserved domains. Again, DWF4 matches the domain C consensus sequence. Finally, the anchoring domain in the N-terminal end was distinguished by a repeat of the hydrophobic residue leucine. In addition, in DWF4, two acidic (glutamate) and two basic (histidine) residues precede the repeated leucine in the N-terminal leader sequence. These charged residues may add more stability to the membrane topology of the protein as a strong start-stop transfer peptide (von Heijne (1988) *Biochim. Biophys. Acta* 947:307-333).

Amend the paragraph beginning at page **55**, line **28**, as follows:

Thus, phylogenetic analyses of these seven proteins with cytochrome P450s unique to plants (group A; Durst and Nelson (1995), *supra*) indicate that DWF4 does not cluster with these cytochrome P450s (Figure 4). Rather, DWF4 clustered with cytochrome P450s from other organisms: cyanobacteria (CYP120), rat (CYP3A2), human (CYP3A3X), and plants (CYP90,

CYP85, and CYP88). DWF4 also deviates from the consensus sequence in the group A heme binding domain in that it possesses a PFGGGPRLCAG (SEQ ID NO:29) sequence in which arginine (R) is substituted for proline (P). However, domain A of DWF4, AGHETS (SEQ ID NO:30), fits the consensus of domain A of group A. These characteristics suggest that DWF4 is a monooxygenase, similar to P450s of group A, that utilizes molecular oxygen as a source of the hydroxyl group, but it mediates some reaction(s) that are not necessarily specific for plants, for instance, steroid hormone biosynthesis, which is a critical event for animals. In fact, the similarity of DWF4 to the rat testosterone 6β-hydroxylase (34%; GenBank accession number 631895) or glucocorticoid-inducible hydroxylase (31%; Molowa et al. 1986; GenBank accession number M13785) supports this idea. Further, the similarity that DWF4 shares with CYP90A and CYP85, 66 and 59%, respectively, is additional proof that it is involved in plant steroid biosynthesis (Bishop et al. 1996 ; Szekeres et al. 1996).